

1600 0420  
0300

TECH CENTER 1600/2900

FEB 20 2002 JUN 20 2002

RECEIVED

645

ENTERED

1645

RAW SEQUENCE LISTING  
 PATENT APPLICATION: US/09/917,378

DATE: 01/29/2002  
 TIME: 12:15:15

Input Set : A:\401977us.app  
 Output Set: N:\CRF3\01292002\I917378.raw

3 <110> APPLICANT: DING, SHI-YOU  
 4 ADNEY, WILLIAM S.  
 5 VINZANT, TODD B.  
 6 DECKER, STEPHEN R.  
 7 HIMMEL, MICHAEL E.  
 9 <120> TITLE OF INVENTION: THERMAL TOLERANT MANNANASE FROM ACIDOTHERMUS  
 10 CELLULOLYTICUS  
 12 <130> FILE REFERENCE: 40197.7US01  
 14 <140> CURRENT APPLICATION NUMBER: 09/917,378  
 15 <141> CURRENT FILING DATE: 2001-07-28  
 17 <160> NUMBER OF SEQ ID NOS: 8  
 19 <170> SOFTWARE: PatentIn Ver. 2.1  
 21 <210> SEQ ID NO: 1  
 22 <211> LENGTH: 762  
 23 <212> TYPE: PRT  
 24 <213> ORGANISM: Acidothermus cellulolyticus  
 26 <400> SEQUENCE: 1  
 27 Met Gly Leu Val Arg Arg Pro Ala Arg Ala Phe Val Ala Thr Ala Ala  
 28 1 5 10 15  
 30 Gly Thr Ala Val Ala Ala Ala Thr Leu Gly Ser Ile Thr Met Pro  
 31 20 25 30  
 33 Ser Ala Thr Ala Ala Pro Ala Gly Phe Val Thr Ala Ser Gly Gly Gln  
 34 35 40 45  
 36 Phe Val Leu Asn Gly Leu Pro Tyr Arg Tyr Gly Gly Thr Asn Asn Tyr  
 37 50 55 60  
 39 Tyr Leu Ser Tyr Gln Ser His Ala Asp Val Asp Asp Val Leu Ala Lys  
 40 65 70 75 80  
 42 Ala Gln Ala Met Asn Leu Ser Val Ile Arg Thr Trp Gly Phe Ile Asp  
 43 85 90 95  
 45 Ile Gly Ser Leu Asp Gly Ser Val Pro Thr Ile Asp Gly Asn Lys Asn  
 46 100 105 110  
 48 Gly Phe Tyr Phe Gln Tyr Trp Asp Pro Ser Thr Gly Ala Pro Ala Tyr  
 49 115 120 125  
 51 Asn Asp Gly Pro Thr Gly Leu Gln Gly Leu Asp Tyr Ala Ile Ala Ser  
 52 130 135 140  
 54 Ala Ala Ala His Gly Leu Arg Val Ile Val Val Leu Thr Asn Asp Trp  
 55 145 150 155 160  
 57 Lys Glu Phe Gly Gly Met Asp Gln Tyr Asp Lys Trp Tyr Gly Leu Pro  
 58 165 170 175  
 60 Tyr His Asp Asn Phe Tyr Thr Asp Pro Arg Thr Gln Gln Ala Tyr Lys  
 61 180 185 190  
 63 Asn Trp Val Asn His Leu Leu Asn Arg Val Asn Ser Ile Thr Gly Val  
 64 195 200 205

## RAW SEQUENCE LISTING

DATE: 01/29/2002

PATENT APPLICATION: US/09/917,378

TIME: 12:15:15

Input Set : A:\401977us.app

Output Set: N:\CRF3\01292002\I917378.raw

TECH CENTER 1600/2900

JUN 20 2002

RECEIVED

```

66 Thr Tyr Lys Asn Asp Pro Thr Ile Phe Ala Trp Glu Leu Ala Asn Glu
67      210                      215                      220
69 Pro Arg Cys Val Gly Ser Gly Thr Leu Pro Thr Ser Gly Thr Cys Thr
70 225                      230                      235                      240
72 Gln Ala Thr Ile Val Asn Trp Val Asp Gln Met Ser Ala Tyr Val Lys
73      245                      250                      255
75 Ser Ile Asp Pro Asn His Met Val Ser Val Gly Asp Glu Gly Phe Tyr
76      260                      265                      270
78 Ile Gly Ser Thr Gln Gly Ser Gly Trp Pro Tyr Asn Asp Pro Ser Asp
79      275                      280                      285
81 Gly Val Asp Asn Asn Ala Leu Leu Arg Val Lys Asn Ile Asp Phe Gly
82      290                      295                      300
84 Thr Tyr His Leu Tyr Pro Asn Tyr Trp Gly Gln Asn Ala Asp Trp Gly
85 305                      310                      315                      320
87 Thr Gln Trp Ile Lys Asp His Ile Ala Asn Ala Ala Ala Ile Gly Lys
88      325                      330                      335
90 Pro Thr Ile Leu Glu Glu Phe Gly Trp Gln Thr Pro Asp Arg Asp Ser
91      340                      345                      350
93 Val Tyr Gln Thr Trp Thr Gln Thr Val Arg Thr Asn Gly Glu Ala Gly
94      355                      360                      365
96 Trp Asn Phe Trp Met Leu Ala Gly Asn Val Asn Gly Gln Pro Tyr Pro
97      370                      375                      380
99 Asn Tyr Asp Gly Phe Asn Val Tyr Tyr Pro Ser Ser Thr Ala Thr Val
100 385                      390                      395                      400
102 Leu Ala Ser Glu Ala Leu Ala Ile Ser Thr Gly Thr Ser Pro Pro Pro
103      405                      410                      415
105 Ser Pro Ser Ser Ser Pro Ser Ser Ser Pro Ser Pro Ser Pro Ser Pro
106      420                      425                      430
108 Ser Ala Ser Pro Ser Ala Ser Pro Ser Ala Ser Ser Ser Pro Ser Pro
109      435                      440                      445
111 Ser Pro Ser Ser Ser Pro Val Ser Gly Gly Val Lys Val Gln Tyr Lys
112      450                      455                      460
114 Asn Asn Asp Ser Ala Pro Gly Asp Asn Gln Ile Lys Pro Gly Leu Gln
115 465                      470                      475                      480
117 Leu Val Asn Thr Gly Ser Ser Ser Val Asp Leu Ser Thr Val Thr Val
118      485                      490                      495
120 Arg Tyr Trp Phe Thr Arg Asp Gly Gly Ser Ser Thr Leu Val Tyr Asn
121      500                      505                      510
123 Cys Asp Trp Ala Ala Met Gly Cys Gly Asn Ile Arg Ala Ser Phe Gly
124      515                      520                      525
126 Ser Val Asn Pro Ala Thr Pro Thr Ala Asp Thr Tyr Leu Gln Leu Ser
127      530                      535                      540
129 Phe Thr Gly Gly Thr Leu Ala Ala Gly Gly Ser Thr Gly Glu Ile Gln
130 545                      550                      555                      560
132 Asn Arg Val Asn Lys Ser Asp Trp Ser Asn Phe Asp Glu Thr Asn Asp
133      565                      570                      575
135 Tyr Ser Tyr Gly Thr Asn Thr Ala Phe Gln Asp Trp Thr Lys Val Thr
136      580                      585                      590
138 Val Tyr Val Asn Gly Arg Leu Val Trp Gly Thr Glu Pro Ser Gly Thr

```

## RAW SEQUENCE LISTING

DATE: 01/29/2002

PATENT APPLICATION: US/09/917,378

TIME: 12:15:15

Input Set : A:\401977us.app

Output Set: N:\CRF3\01292002\I917378.raw

```

139          595          600          605
141 Ser Pro Ser Pro Thr Pro Ser Pro Ser Pro Thr Pro Ser Pro Ser Pro
142          610          615          620
144 Ser Pro Thr Pro Ser Pro Ser Ser Ser Pro Ser Pro Ser Pro Ser Pro
145 625          630          635          640
147 Ser Pro Ser Pro Thr Pro Ser Pro Ser Pro Ser Pro Ser Pro Ser Pro
148          645          650          655
150 Ser Val Ser Ser Ser Gly Val Gly Cys Arg Ala Thr Tyr Val Val Asn
151          660          665          670
153 Ser Asp Trp Gly Ser Gly Phe Thr Ala Thr Val Thr Val Thr Asn Thr
154          675          680          685
156 Gly Ser Arg Ala Thr Ser Gly Trp Thr Val Ala Trp Ser Phe Gly Gly
157          690          695          700
159 Asn Gln Thr Val Thr Asn Tyr Trp Asn Thr Ala Leu Thr Gln Ser Gly
160 705          710          715          720
162 Ala Ser Val Thr Ala Thr Asn Leu Ser Tyr Asn Asn Val Ile Gln Pro
163          725          730          735
165 Gly Gln Ser Thr Thr Phe Gly Phe Asn Gly Ser Tyr Ser Gly Thr Asn
166          740          745          750
168 Thr Ala Pro Thr Leu Thr Cys Thr Ala Ser
169          755          760
172 <210> SEQ ID NO: 2
173 <211> LENGTH: 2289
174 <212> TYPE: DNA
175 <213> ORGANISM: Acidothermus cellulolyticus
177 <400> SEQUENCE: 2
178 atgggtctag tgcgtcgccc tgcgcgagca tttgttgcca ccgcggcccg cactgccgtt 60
179 gctgccgcgg cgacgctcgg ctcaatcacc atgccgtcag ccacggcagc gccggcgcca 120
180 ttcgtcaccg catccggcgg tcagttcggt ctgaacggcc ttccctatcg ttacggggga 180
181 acgaacaact attacctcag ctatcagtcg caccgcgacg tcgatgacgt gttggccaag 240
182 gctcaagcga tgaatctttc tgtcatccgg acctgggggt tcatcgacat cggctctctt 300
183 gacggctccg tgcccacaat cgatggcaac aagaacggct tctactttca gtactgggac 360
184 ccgtcgaccg gcgctccggc gtacaacgac gggccgaccg gcttgcaagg ccttgactac 420
185 gcgatcgcca gcgcggccgc gcaaggcctt cgggtgattg tcgtcctcac caacgactgg 480
186 aaagaatttg ggggaatgga tcaatacgac aagtggtagc gccttcctta ccacgacaac 540
187 ttctacaccg acccccggac ccagcaggcg tacaagaatt gggatcaatc tctactgaac 600
188 cgggtcaaca gcattaccgg cgtgacgtac aagaacgata caacgatctt tgcttgggaa 660
189 cttgccaatg agccgcgctg cgtaggaagc ggcacattac caacctcggg cactgtcact 720
190 caggcgacca ttgtcaactg ggtcgatcaa atgtcggcgt acgtcaaaaag catagaccct 780
191 aaccatatgg tctcggtcgg cgacgaaggg ttctacattg ggtcaacgca ggggaagcggc 840
192 tggccataca acgacccgtc cgacggcgtc gacaacaatg ctcttctccg tgtcaagaac 900
193 attgactttg gcacgtatca cctgtacccg aattactggg gccagaacgc ggactggggg 960
194 acgcaatgga tcaaggatca tattgcgaat gccgcagcga tcggcaagcc gaccattctc 1020
195 gaagaattcg gctggcagac accggaccgc gattccgtct atcagacgtg gaccagact 1080
196 gtgcgtacga acggtgaagc aggttggaac ttctggatgc tcgctgggaa tgtcaacggc 1140
197 cagccatata cgaactatga cggttcaac gtctactacc caagttcaac agcgaccgtc 1200
198 ctgccagcg aggcgctcgc aatcagtacc ggcacatcgc ctccgcgctc gccgagctcg 1260
199 agtccatcct cgtcgccgtc tccgtcgccg tctccgtcgg cgtctccgtc ggcgtctccg 1320
200 tcggcgctct cgtcgccgag cccgtctccg tcgtcgctcg cgggtgtcggg tgggggtgaag 1380

```

## RAW SEQUENCE LISTING

DATE: 01/29/2002

PATENT APPLICATION: US/09/917,378

TIME: 12:15:15

Input Set : A:\401977us.app

Output Set: N:\CRF3\01292002\I917378.raw

```

201 gtgcagtaca agaacaatga ttcggcgccg ggtgataacc agatcaaacc ggggtctccag 1440
202 ttggtgaata cgggggtcgtc gtcggtggat ttgtcgacgg tgacgggtgc gtactgggttc 1500
203 acccgggatg gtgggtcgtc gacactgggtg tacaactgtg actggggcggc gatgggggtgt 1560
204 gggaatatcc gcgcctcgtt cggctcgggtg aaccggcgga cgccgacggc ggacacctac 1620
205 ctgcagttgt cgttctactgg tggaaacgttg gccgctgggtg ggtcgacggg tgagattcaa 1680
206 aaccgggtga ataagagtga ctggtcgaac tttgatgaga ccaatgacta ctcgtatggg 1740
207 acgaacaccg ccttccagga ttggacgaag gtgacgggtg atgtcaatgg ccggtcgggtg 1800
208 tgggggactg aaccgtccgg caccagcccc agccccacac ccagccccag cccaaccccc 1860
209 tccccgagcc cgagcccgac cccaagcccc agctcctccc catccccgtc cccgagcccc 1920
210 agccccagcc ctacgccgtc cccgtcggcg agcccgctgc cgtcgccgag tgtgtcgtcg 1980
211 tcgggtgttg ggtgccgggc gacgtatgtg gtgaatagtg attgggggttc tgggtttacg 2040
212 gcgacgggtg cgggtgacgaa taccgggagc cgggcgacga gcgggtggac ggtggcgtgg 2100
213 tcgtttgggtg ggaatcagac ggtcacgaac tactggaaca ctgcgttgac ccaatcaggt 2160
214 gcatcgggtg cggcgacgaa cctgagctac aacaacgtga tccaacgggg tcagtcgacc 2220
215 accttcggat tcaacggaag ttactcagga acaaacaccg cacctacact cacctgcacg 2280
216 gctagttga 2289
219 <210> SEQ ID NO: 3
220 <211> LENGTH: 375
221 <212> TYPE: PRT
222 <213> ORGANISM: Artificial Sequence
224 <220> FEATURE:
225 <223> OTHER INFORMATION: Description of Artificial Sequence: GH5 catalytic
226 domain
228 <400> SEQUENCE: 3
229 Ala Pro Ala Gly Phe Val Thr Ala Ser Gly Gly Gln Phe Val Leu Asn
230 1 5 10 15
232 Gly Leu Pro Tyr Arg Tyr Gly Gly Thr Asn Asn Tyr Tyr Leu Ser Tyr
233 20 25 30
235 Gln Ser His Ala Asp Val Asp Asp Val Leu Ala Lys Ala Gln Ala Met
236 35 40 45
238 Asn Leu Ser Val Ile Arg Thr Trp Gly Phe Ile Asp Ile Gly Ser Leu
239 50 55 60
241 Asp Gly Ser Val Pro Thr Ile Asp Gly Asn Lys Asn Gly Phe Tyr Phe
242 65 70 75 80
244 Gln Tyr Trp Asp Pro Ser Thr Gly Ala Pro Ala Tyr Asn Asp Gly Pro
245 85 90 95
247 Thr Gly Leu Gln Gly Leu Asp Tyr Ala Ile Ala Ser Ala Ala Ala His
248 100 105 110
250 Gly Leu Arg Val Ile Val Val Leu Thr Asn Asp Trp Lys Glu Phe Gly
251 115 120 125
253 Gly Met Asp Gln Tyr Asp Lys Trp Tyr Gly Leu Pro Tyr His Asp Asn
254 130 135 140
256 Phe Tyr Thr Asp Pro Arg Thr Gln Gln Ala Tyr Lys Asn Trp Val Asn
257 145 150 155 160
259 His Leu Leu Asn Arg Val Asn Ser Ile Thr Gly Val Thr Tyr Lys Asn
260 165 170 175
262 Asp Pro Thr Ile Phe Ala Trp Glu Leu Ala Asn Glu Pro Arg Cys Val
263 180 185 190
265 Gly Ser Gly Thr Leu Pro Thr Ser Gly Thr Cys Thr Gln Ala Thr Ile

```

## RAW SEQUENCE LISTING

DATE: 01/29/2002

PATENT APPLICATION: US/09/917,378

TIME: 12:15:15

Input Set : A:\401977us.app

Output Set: N:\CRF3\01292002\I917378.raw

```

266          195          200          205
268 Val Asn Trp Val Asp Gln Met Ser Ala Tyr Val Lys Ser Ile Asp Pro
269          210          215          220
271 Asn His Met Val Ser Val Gly Asp Glu Gly Phe Tyr Ile Gly Ser Thr
272 225          230          235          240
274 Gln Gly Ser Gly Trp Pro Tyr Asn Asp Pro Ser Asp Gly Val Asp Asn
275          245          250          255
277 Asn Ala Leu Leu Arg Val Lys Asn Ile Asp Phe Gly Thr Tyr His Leu
278          260          265          270
280 Tyr Pro Asn Tyr Trp Gly Gln Asn Ala Asp Trp Gly Thr Gln Trp Ile
281          275          280          285
283 Lys Asp His Ile Ala Asn Ala Ala Ala Ile Gly Lys Pro Thr Ile Leu
284          290          295          300
286 Glu Glu Phe Gly Trp Gln Thr Pro Asp Arg Asp Ser Val Tyr Gln Thr
287 305          310          315          320
289 Trp Thr Gln Thr Val Arg Thr Asn Gly Glu Ala Gly Trp Asn Phe Trp
290          325          330          335
292 Met Leu Ala Gly Asn Val Asn Gly Gln Pro Tyr Pro Asn Tyr Asp Gly
293          340          345          350
295 Phe Asn Val Tyr Tyr Pro Ser Ser Thr Ala Thr Val Leu Ala Ser Glu
296          355          360          365
298 Ala Leu Ala Ile Ser Thr Gly
299          370          375
302 <210> SEQ ID NO: 4
303 <211> LENGTH: 154
304 <212> TYPE: PRT
305 <213> ORGANISM: Artificial Sequence
307 <220> FEATURE:
308 <223> OTHER INFORMATION: Description of Artificial Sequence: Carbohydrate
309 binding domain
311 <400> SEQUENCE: 4
312 Val Ser Gly Gly Val Lys Val Gln Tyr Lys Asn Asn Asp Ser Ala Pro
313 1          5          10          15
315 Gly Asp Asn Gln Ile Lys Pro Gly Leu Gln Leu Val Asn Thr Gly Ser
316          20          25          30
318 Ser Ser Val Asp Leu Ser Thr Val Thr Val Arg Tyr Trp Phe Thr Arg
319          35          40          45
321 Asp Gly Gly Ser Ser Thr Leu Val Tyr Asn Cys Asp Trp Ala Ala Met
322          50          55          60
324 Gly Cys Gly Asn Ile Arg Ala Ser Phe Gly Ser Val Asn Pro Ala Thr
325 65          70          75          80
327 Pro Thr Ala Asp Thr Tyr Leu Gln Leu Ser Phe Thr Gly Gly Thr Leu
328          85          90          95
330 Ala Ala Gly Gly Ser Thr Gly Glu Ile Gln Asn Arg Val Asn Lys Ser
331          100          105          110
333 Asp Trp Ser Asn Phe Asp Glu Thr Asn Asp Tyr Ser Tyr Gly Thr Asn
334          115          120          125
336 Thr Ala Phe Gln Asp Trp Thr Lys Val Thr Val Tyr Val Asn Gly Arg
337          130          135          140

```

VERIFICATION SUMMARY

DATE: 01/29/2002

PATENT APPLICATION: US/09/917,378

TIME: 12:15:16

Input Set : A:\401977us.app

Output Set: N:\CRF3\01292002\I917378.raw